On fixing semantic alignment evaluation measures

Jérôme David, Jérôme Euzenat

On fixing semantic alignment evaluation measures

Jérôme David and Jérôme Euzenat

LIG & INRIA Grenoble Rhône-Alpes
Grenoble, France
Jerome.{David|Euzenat}@inrialpes.fr

Abstract. The evaluation of ontology matching algorithms mainly consists of comparing a produced alignment with a reference one. Usually, this evaluation relies on the classical precision and recall measures. This evaluation model is not satisfactory since it does not take into account neither the closeness of correspondances, nor the semantics of alignments. A first solution consists of generalizing the precision and recall measures in order to solve the problem of rigidity of classical model. Another solution aims at taking advantage of the semantic of alignments in the evaluation. In this paper, we show and analyze the limits of these evaluation models. Given that measures values depend on the syntactic form of the alignment, we first propose an normalization of alignment. Then, we propose two new sets of evaluation measures. The first one is a semantic extension of relaxed precision and recall. The second one consists of bounding the alignment space to make ideal semantic precision and recall applicable.

1 Introduction

With the semantic Web, many related but heterogenous ontologies are being created. In such an open context, there is no reason why two domain-related applications would share the same ontologies. In order to facilitate the exchange of knowledge between such applications, ontology matching aims at discovering a set of relations between entities from two ontologies. This set of relations is called an alignment.

Many different matching algorithms have been designed [Euzenat and Shvaiko, 2007]. In order to compare the performance of such algorithms, some efforts are devoted to the evaluation of ontology matching tools. Since 2004, the Ontology Alignment Evaluation Initiative[1] (OAEI) organizes, every year, an evaluation of ontology matching methods. The evaluation of matching algorithms consists of comparing a produced alignment with a reference one. This evaluation often relies on two classical measures used in information retrieval: precision and recall [van Rijsbergen, 1979]. Precision measures the ratio of correct correspondences in the evaluated alignment. Recall measures the ratio of reference correspondence found by the evaluated alignment.

In the context of alignment evaluation, precision and recall present the drawbacks to be all-or-nothing measures [Ehrig and Euzenat, 2005] and they do not consider neither the semantic of alignment relations, nor those of ontologies. Then, an alignment can be very close to the expected result and have low precision and recall values. Two approaches have been proposed for correcting these drawbacks. [Ehrig and Euzenat,

[1] http://oaei.ontologymatching.org

2005] introduced a generalization of precision and recall measures. This approach relies on syntactic measures relaxing the all-or-nothing feature of classical measures in order to take into account close correspondences. [Euzenat, 2007] has introduced semantic precision and recall measures which rely on a semantic of alignments

We will show that semantic precision and recall are still dependent on the alignment syntax and, as a consequence, they can assign different values to semantically equivalent alignments. [David, 2007] proposes to use the ideal semantic precision and recall measures introduced in [Euzenat, 2007] restricted to alignments containing only simple correspondences, i.e. only between named entities.

In this paper, we show and analyze the limits and problems of the semantic precision and recall measures. To overcome their drawbacks, we first investigate an approach allowing to normalize alignments. This normalization relies on algebra of alignment relations and can partially resolves problems encountered by the evaluation measures. In addition, we propose two adaptations of the relaxed and semantic measures. The first adaptation makes use of the generalization framework of [Ehrig and Euzenat, 2005] and allows to locally consider the semantic of alignments. The second one is a restriction of Semantic closures of alignment. This restriction makes the ideal semantic measures proposed in [Euzenat, 2007] useable.

This paper is organized as follows: a first section introduces the definitions related to the syntax and semantics of alignments. In the second section, we first present and introduce five properties that an ideal model should satisfy. Then, we present the classical evaluation measures and the semantic evaluation measures which satisfy three of the five desired properties. In the following section, we explain why these semantic measures do not satisfy the two last properties. The last section proposes three ways for fixing the semantic measures: a normalization of alignments, new relaxed semantic precision and recall measures, and Λ-bounded semantic evaluation measures.

2 Ontology alignment: syntax and semantic

2.1 Definition and syntax

An alignment groups correspondences between entities or formulas from two ontologies o_1 and o_2. Each element of correspondence can be associated to a quality value by a function q. We use the following syntax for representing an alignment:

Definition 1 (Alignment). *An **alignment** between two ontologies o_1 and o_2 is a set of correspondances holding between o_1 and o_2. A correspondance, noted $c = (x, y, \mathcal{R})$, is a triple where x, respectively y, are formulas (or entities) from o_1, respectively from o_2, and $\mathcal{R}$ is the relation holding between x and y. A correspondance $c = (x, y, \mathcal{R})$ can be also written $x\mathcal{R}y$*

This definition includes both simple alignments considering only matching relations between entities (classes or properties) and complex alignments containing relations between formula inferred from the ontologies.

2.2 Semantic of alignments: Semantic closure and semantic reduction

Alignments between ontologies can be helpful for reasoning with several ontologies. For enabling reasoning capabilities, a semantic for alignments must be defined. In this paper, we relies on the semantic proposed in [Euzenat, 2007]. This semantic of alignment is function of the semantics of each individual ontology. The semantic of an ontology is given by its set of models.

Definition 2 (Model). *a model $m = \langle I, D \rangle$ of o is a function I from the terms of o to a domain of interpretation D, which satisfies all the assertions in o:*

$$\forall \delta \in o, m \models \delta$$

The set of models of an ontology o is denoted as $\mathcal{M}(o)$.

Because the models of various ontologies can have different interpretation domains, we use the notion of an equalising function, which helps make these domains commensurate.

Definition 3 (Equilising function). *Given a family of interpretations $\langle I_o, D_o \rangle_{o \in \Omega}$ of a set of ontologies Ω, an equalising function for $\langle I_o, D_o \rangle_{o \in \Omega}$ is a family of functions $\gamma = (\gamma_o : D_o \longrightarrow U)_{o \in \Omega}$ from the ontology domains of interpretation to a global domain of interpretation U. The set of all equalising functions is called Γ.*

The relations used in correspondences do not necessarily belong to the ontology languages. As a consequence, a semantics for them must be provided.

Definition 4 (Interpretation of alignment relations). *Given $\mathcal{R}$ an alignment relation and U a global domain of interpretation, $\mathcal{R}$ is interpreted as a binary relation over U, i.e., $\mathcal{R}^U \subseteq U \times U$.*

The definition of correspondence satisfiability relies on γ and the interpretation of relations. It requires that in the equalised models, the correspondences are satisfied.

Definition 5 (Satisfied correspondence). *A correspondence $c = \langle x, y, \mathcal{R} \rangle$ is satisfied for an equalising function γ by two models m, m' of o, o' if and only if $\gamma_o \cdot m \in \mathcal{M}(o)$, $\gamma_{o'} \cdot m' \in \mathcal{M}(o')$ and*

$$\langle \gamma_o(m(e)), \gamma_{o'}(m'(e')) \rangle \in \mathcal{R}^U$$

This is denoted as $m, m' \models_\gamma c$.

Given an alignment between two ontologies, the semantics of the aligned ontologies can be defined as follows.

Definition 6 (Models of aligned ontologies). *Given two ontologies o and o' and an alignment A between these ontologies, a model m'' of these ontologies aligned by A is a triple $\langle m, m', \gamma \rangle \in \mathcal{M}(o) \times \mathcal{M}(o') \times \Gamma$, such that $m, m' \models_\gamma A$.*

We will consider a specific kind of consequence, α-consequences [Euzenat, 2007], which are the correspondences holding for all models of aligned ontologies.

Definition 7 **(α-Consequence of aligned ontologies).** *Given two ontologies o and o' and an alignment A between these ontologies, a correspondence δ is a α-consequence of o, o' and A (noted $A \models \delta$) if and only if for all models $\langle m, m', \gamma \rangle$ of o, o' and A, $m, m' \models_\gamma \delta$ (the set of α-consequences is noted by $Cn(A)$).*

Given this semantic, the semantic closure and semantic reduction of an alignment are given by the following definitions:

Definition 8 **(Semantic closure).** *The semantic closure $Cn(A)$ of an alignment A is the set of its α-consequences.*

Obviously , the semantic closure of an alignment is unique but it has no reason to be finite.

Definition 9 **(Semantic reduction).** *A semantic reduction (or minimal cover) A^0 of an alignment A is an alignment satisfying $Cn(A^0) = Cn(A)$ and $\forall c \in A^0,\ Cn(A^0 - \{c\}) \neq Cn(A)$*

There could exist several semantic reductions for a given alignment. An alignment A contains redundant elements if A is not a minimal cover. A correspondence $c \in A$ is redundant if $A - \{c\} \models c$.

3 Evaluation models

Alignment evaluation is achieved by comparing the produced alignment with the reference one. This comparison usually relies on the precision (P) and the recall (R) measures [van Rijsbergen, 1979]. Intuitively, the precision aims at measuring the correctness of the evaluated alignment. The recall is used for quantifying the completeness of the evaluated alignment.

In the rest of this paper, we will consider two alignments between ontologies o_1 and o_2: a reference alignment, noted A_r, and an alignment produced by some matching method A_e.

3.1 Desired properties of evaluation measures

If we consider that precision and recall should approximate correctness and completeness, an ideal model taking semantic into account, would respect the constraints given by [Euzenat, 2007]:

- $A_r \models A_e \Rightarrow P(A_e, A_r) = 1$ (max-correctness)
- $A_e \models A_r \Rightarrow R(A_e, A_r) = 1$ (max-completeness)
- $Cn(A_e) = Cn(A_r)$ iff $P(A_e, A_r) = 1$ and $R(A_e, A_r) = 1$ (definiteness)

Furthermore, in the evaluation context, one could be interested to compare several alignments produced by some matching algorithms against one reference alignment. Then, it would be useful that two semantically equivalent alignments have the same precision and recall values.

– $Cn(A_{e_1}) = Cn(A_{e_2}) \Rightarrow P(A_{e_1}, A_r) = P(A_{e_2}, A_r)$ and $R(A_{e_1}, A_r) = R(A_{e_2}, A_r)$ (semantic-equality)

Finally, if the evaluated and the reference alignments share some common information then the precision and recall values must not be null:

– $P(A_e, A_r) = 0$ and $R(A_e, A_r) = 0$ iff $Cn(A_e) \cap Cn(A_r) = Cn(\emptyset)$ (overlapping positiveness)

3.2 Classical evaluation model

The classical evaluation model is based on the interpretation of alignments as sets and considers the following sets :

– ***E***: the set of all correspondences that could be generated between o_1 and o_2. This set is a subset of the cartesian product of all entities which can be deduced from o_1, those deductible for o_2 and the set of matching relations considered.
– **true-positives**: the set of correspondences which are found by the matching method and contained in the reference alignment.
– **false-positives**: the set of correspondences which are found by the matching method but not contained in the reference alignment.
– **false-negatives**: the set of reference correspondences which are not found by the matching method.
– **true-negatives**: the set of correspondences that are neither in the evaluated alignment nor in the reference alignment.

	relevant	not relevant	
found	$\lvert A_e \cap A_r \rvert$ true-positives	$\lvert A_e - A_r \rvert$ false-positives	$\lvert A_e \rvert$
not found	$\lvert A_r - A_e \rvert$ false-negatives	$\lvert (E - A_r) - A_e \rvert$ true-negatives	$\lvert E - A_e \rvert$
	$\lvert A_r \rvert$	$\lvert E - A_r \rvert$	

Table 1. Contingency of sets A_e and A_r.

The cardinalities of these sets are given in the contingency table 1. The sets of true-positives, false negatives, and false-positives are defined only from A_e and A_r. The set of true-negatives is also function of the set E which is not easily identifiable.

From these contingencies, the classical measure of precision and recall can be defined. The precision (P) represents the proportion of found correspondences that are relevant:

$$P(A_e, A_r) = \frac{|A_e \cap A_r|}{|A_e|} \tag{1}$$

The recall (R) represents the proportion of relevant correspondences that have been found :

$$R(A_e, A_r) = \frac{|A_e \cap A_r|}{|A_r|} \tag{2}$$

3.3 Limitations of classical precision and recall

These two measures applied to this simple model have the advantages to be easily computable and understandable. However, they verify none of the constraints presented Section 3.1. This is because they do not consider the semantic of alignment relations, nor the semantic of ontologies.

Firstly, they do not take into account the semantic of matching relations. For example, if the produced alignment A_e contains the elements $x \sqsubseteq y$ and $x \sqsupseteq y$, and the reference alignment A_r contains the element $x \equiv y$, then the classical model will consider $x \sqsubseteq y$ and $x \sqsupseteq y$ as false-positives and $x \equiv y$ as false-negative. In this case the precision and recall values are equals to 0 even if $A_e \equiv A_r$.

Secondly, this classical model does not take the semantic of ontologies into account. For example, A_e contains the element $x' \sqsubseteq y$, the reference alignment A_r contains the element $x \equiv y$, and the ontology o_1 states $x' \sqsubseteq x$. Even if $A_r \models A_e$, the classical precision will be equal to 0 since the correspondence $x' \sqsubseteq y$ is considered as a false-positive by this evaluation model.

3.4 Semantic evaluation models

In order to resolve the drawbacks of classical precision and recall, [Euzenat, 2007] proposes to take into account the semantics of matching relations and ontologies. The author provides two extensions of precision and recall.

The ideal extension of the classical model consists of replacing A_e and A_r by their respective sets of α-consequences, $Cn(A_e)$ and $Cn(A_r)$. Table 2 show the new contingencies.

	relevant	not relevant	
found	$\|Cn(A_e) \cap Cn(A_r)\|$ true-positives	$\|Cn(A_e) - Cn(A_r)\|$ false-positives	$\|Cn(A_e)\|$
not found	$\|Cn(A_r) - Cn(A_e)\|$ false-negatives	$\|(E - Cn(A_r)) - Cn(A_e)\|$ true-negatives	$\|E - Cn(A_e)\|$
	$\|cn(A_r)\|$	$\|E - Cn(A_r)\|$	

Table 2. Contingencies of the ideal extension of the classical model.

From this extended model, ideal precision and recall measures, respectively named P_i and R_i, are :

$$P_i(A_e, A_r) = \frac{|Cn(A_e) \cap Cn(A_r)|}{|Cn(A_e)|} \tag{3}$$

$$R_i(A_e, A_r) = \frac{|Cn(A_e) \cap Cn(A_r)|}{|Cn(A_r)|} \tag{4}$$

These measures correct the drawbacks of the classical model and all the properties given Section 3.1 are satisfied. However, they bring a new problem : as the semantic closures of alignments could be infinite, then the measures may be undefined.

In order to overcome this problem, [Euzenat, 2007] introduces two new measures known as semantic precision and semantic recall.

Semantic precision measures the proportion of evaluated correspondances of A_e that can be deduced from A_r.

$$P_s(A_e, A_r) = \frac{|A_e \cap Cn(A_r)|}{|A_e|} \tag{5}$$

Semantic recall measures the proportion of reference correspondances of A_r that can be deduced from A_e.

$$R_s(A_e, A_r) = \frac{|Cn(A_e) \cap A_r|}{|A_r|} \tag{6}$$

With these measures, the max-correctness, max-completeness, and definiteness properties are preserved. The values of semantic precision and semantic recall are greater than or equal to those of classical ones because $|A_e \cap Cn(A_r)| > |A_e \cap A_r|$ and $|Cn(A_e) \cap A_r| > |A_e \cap A_r|$.

4 Limitations of semantic precision and recall

Semantic precision and recall correct some drawback of classical precision and recall measure since they satisfy the max-correctness, max-completeness, and definiteness properties. Nevertheless, they do not satisfy the semantic-equality and overlapping-positiveness properties which we have introduced. This is due to the fact that these semantic measures are still dependent on the syntactic form of the alignments.

4.1 Limitation concerning semantic-equality property

Two alignments A_{e_1} and A_{e_2} having the same closure and then, semantically equivalent, could have different precision and recall values according to A_r. This due to the fact that the semantic precision and recall are directly function of the cardinalities of the correspondences sets which could be different for two semantically equivalent alignments.

We give two examples demonstrating that semantic evaluation measures do not satisfy the semantic-equality property. In the first example, we reason only with alignment. In the second example, we show that redundancy in alignment can break the satisfaction of semantic-equality property by precision measure.

In the first example, we consider two alignments $A_{e_1} = \{x \equiv y, u \equiv v\}$ and $A_{e_2} = \{x \sqsubseteq y, x \sqsupseteq y, u \equiv v\}$. These two alignments are equivalent since we have only replaced the equivalence $x \equiv y$ of A_{e_1} by $x \sqsubseteq y$ and $x \sqsupseteq y$ in A_{e_2}. According to a reference alignment $A_r = \{x \equiv y\}$, the two alignments do not have the same precision values: $P_s(A_{e_1}, A_r) = 1/2$ and $P_s(A_{e_2}, A_r) = 2/3$.

In the second example, we now have the alignments $A_{e_1} = \{x \equiv y, u \equiv v\}$ and $A_{e_2} = \{x' \sqsubseteq y, x \equiv y, u \equiv v\}$, and the knowledge $o_1 \models x' \sqsubseteq x$. These two alignments are equivalent since $x' \sqsubseteq y$ is redundant according to $x \equiv y$. Nevertheless, the semantic precision values will be different: $P_s(A_{e_1}, A_r) = 1/2$ and $P_s(A_{e_2}, A_r) = 2/3$

4.2 Limitation concerning overlapping-positiveness property

An alignment could have a null precision or/and recall value even if the intersection of its consequence sets and those of the reference is not the empty set. This due to the fact that the semantic precision and recall partially take the alignment semantic into account. A correspondance can entail several correspondances. Such a correspondance can be partially true-positive in the sense that it entails a true-positive element but also a false-negative or false-positive element. With the semantic precision and recall, such elements are entirely considered as false-positives or/and false-negatives.

For example, let be the two alignments $A_e = \{a\}$ and $A_r = \{b\}$, another matching relation c and the properties $a \models c$, $b \models c$, $a \not\models b$ and $b \not\models a$. On this trivial example, the semantic precision and recall values are both equals to 0 even if the intersection of their Semantic closures is not equals to the empty set (i.e. $c \in Cn(A_e) \cap Cn(A_r)$).

5 Corrections of semantic evaluation measures

In previous section, we highlighted some drawbacks of classical precision and recall and semantic precision and recall. The first kind of problems concerns the inability of classical and generalized precision and recall measures to reason with the alignment relations. The semantic precision and recall try to resolve this problem by using Semantic closures, but these measures are still defined on the alignment cardinality which is dependent on the syntactic form of the alignment. As a consequence, there are some cases where the semantic-equality property is not satisfied.

When this problem is entirely due to the syntactic form the alignments, we may try to resolve it by normalizing the alignment representation. We propose here a normalization strategy which relies on algebras of alignment relations [Euzenat, 2008].

Then, with the help of alignment normalization, we propose two new sets of evaluation measures. The first one concerns relaxed semantic measures based on the generalized precision and recall framework of [Ehrig and Euzenat, 2005]. Contrarily to the original generalized precision and recall measures provided in the aforementioned paper, these new measures are not only based on the syntactic form of alignment, but also on the semantic of alignments.

The second set of measures is an adaptation of ideal semantic measures of [Euzenat, 2007].

5.1 Normalization of alignments

For allowing measures to respect the semantic-equality property, it is useful to introduce a notion of a normal form for alignments. A normal form for alignments ensures that two semantically equivalent alignments have always the same syntax or form. Naturally, it is a very difficult problem but we can propose a partial solution which only considers the semantic of alignment relations. Our notion of normal form takes benefit of entailment capabilities provided by algebras of alignment relations and does not use any knowledge about the aligned ontologies.

An algebra of alignment relations [Euzenat, 2008] is a particular type of relation algebra [Tarski, 1941] defined by the tuple $\langle 2^{\Gamma}, \cap, \cup, \cdot, \Gamma, \emptyset, \{\equiv\}, ^{-1} \rangle$ where Γ is the set of all elementary relations; $\cap$ and $\cup$ are set-operations used to meet and join two sets of relations, for example, if $x\mathcal{R}y$ or $x\mathcal{R}'y$ then $x\mathcal{R}\cup\mathcal{R}'y$; $\cdot$ is the composition operator, i.e. an associative internal composition law with $\{\equiv\}$ as unity element; $^{-1}$ the converse operator. For instance, if $\Gamma = \{\sqsubset, \sqsupset, \equiv, \between, \bot\}$, all all elementary relations, except $\sqsubset$ and $\sqsupset$, are there own converse and, $\sqsubset^{-1}{=}\sqsupset$ and $\sqsupset^{-1}{=}\sqsubset$.

Such an algebra allows to write any relation between entities (or formulas) as a disjunction of elementary relations. For example, $x \sqsubseteq y$ would be written $x\{\sqsubset, \equiv\}y$. With the help of this relation algebra, any pair of entities or formulas will appear at most once in the alignment.

Definition 10. *An **alignment in normal form** is an alignment $A = (V, q)$ where the set of correspondances V satisfies the following properties:*

1. *$V \subset \{x\mathcal{R}y | x \in o_1 \wedge y \in o_2 \wedge\ \mathcal{R} \subseteq \Gamma\}$: all relations between two entities (or formulas) are written with a disjunction of elementary relations.*
2. *$\forall x\mathcal{R}y \in V,\ \nexists x\mathcal{R}'y \in V,\ \mathcal{R} = \mathcal{R}'$: any pair of entities (or formulas) appear at most once in the alignment.*

Using such a normalization allows to correct classical and semantic precision and recall when relations between entities or formulas are split into several correspondances. For example, let be $A_e = \{x \sqsubseteq y, x \sqsupseteq y\}$ and $A_r = \{x \equiv y\}$. By rewriting these alignments using disjunction of elementary relations, $A_e = \{x\{\sqsubset, \equiv\} \cap \{\sqsupset, \equiv\}y\} = \{x\{\equiv\}y\}$ and $A_r = \{x\{\equiv\}y\}$ will be syntactically equivalent.

Of course, when this problem is due to the semantic of ontologies such a normalization is not sufficient. For example, let be $A_e = \{x \sqsubseteq y\}$ and $A_r = \{x \sqsubseteq z\}$ and the axiom $y \equiv z \in o_2$. These alignments are equivalent (given the previous axiom), but their normalization ($A_e = \{x\{\sqsubset, \equiv\}y\}$ and $A_r = \{x\{\sqsubset, \equiv\}z\}$) are not equal.

5.2 Relaxed semantic precision and recall

In generalized precision and recall framework, evaluation measures are function of a measure quantifying the proximity between two correspondences [Ehrig and Euzenat, 2005]. We propose new proximity measures σ dealing partially with the semantic of alignments. We want such measures to locally respect the max-correctness and max-completeness properties contrarily to those provided in [Ehrig and Euzenat, 2005]:

- if $x'\mathcal{R}'y' \models x\mathcal{R}y$ then $\sigma_{prec}(x\mathcal{R}y, x'\mathcal{R}'y') = 1$ (local max-correctness)
- if $x\mathcal{R}y \models x'\mathcal{R}'y'$ then $\sigma_{rec}(x\mathcal{R}y, x'\mathcal{R}'y') = 1$ (local max-completeness)

In order to propose such measures, we suggest to take advantage of an algebra of alignment relations as presented in the previous section (Section 5.1). Following the example of the relaxed precision and recall, which are oriented, we introduce two new σ measures: σ_{prec} for precision and σ_{rec} for the recall. In these two σ measures, we do not consider the confidence values.

In a first instance, we only consider the case where we have two correspondances aligning the same entities or formulas. Let $x\mathcal{R}y$ and $x\mathcal{R}'y$ be two such correspondances. From the algebra of alignment relations, we have the following properties:

- if $x\mathcal{R}'y \models x\mathcal{R}y$, then $\mathcal{R}' \subseteq \mathcal{R}$,
- if $x\mathcal{R}y \models x\mathcal{R}'y$, then $\mathcal{R} \subseteq \mathcal{R}'$.

Hence, σ_{prec} and σ_{rec} are defined by:

$$\sigma_{prec}(\mathcal{R}, \mathcal{R}') = \frac{|\mathcal{R} \cap \mathcal{R}'|}{|\mathcal{R}'|} \tag{7}$$

$$\sigma_{rec}(\mathcal{R}, \mathcal{R}') = \frac{|\mathcal{R} \cap \mathcal{R}'|}{|\mathcal{R}|} \tag{8}$$

Now, for extending these measures to correspondances which do not align the same entities or formulas, we propose to use this relation algebra also with the ontologies.

Definition 11 (Relaxed semantic proximity measures). *Given an evaluated relation $x\mathcal{R}y$, a reference relation $x'\mathcal{R}'y'$, and relations deduced from ontologies, $o_1 \models x\mathcal{R}_1x'$ and $o_2 \models y\mathcal{R}_2y'$, the relaxed semantic proximities σ_{prec} and σ_{rec} are defined by:*

$$\sigma_{prec}(x\mathcal{R}y, x'\mathcal{R}'y') = \frac{|\mathcal{R} \cap \left(\mathcal{R}_1 \cdot \mathcal{R}' \cdot \mathcal{R}_2^{-1}\right)|}{|\mathcal{R}_1 \cdot \mathcal{R}' \cdot \mathcal{R}_2^{-1}|} \tag{9}$$

$$\sigma_{rec}(x\mathcal{R}y, x'\mathcal{R}'y') = \frac{|(\mathcal{R}_1^{-1} \cdot \mathcal{R} \cdot \mathcal{R}_2) \cap \mathcal{R}'|}{|\mathcal{R}_1^{-1} \cdot \mathcal{R} \cdot \mathcal{R}_2|} \tag{10}$$

The relaxed semantic proximity measures satisfy the local max-correctness and local max-completeness properties. As a consequence, they allow to provide relaxed semantic precision and recall measures which partially deals with alignment semantics. However, such measures do not consider the whole alignment semantic and then, they do not necessarily satisfy any property mentioned Section 3.1.

In our opinion, these semantic proximity measures are a first step for providing new semantic evaluations measures satisfying the desired properties. However, for satisfying these properties, it would be essential to propose new generalized precision and recall measures.

5.3 Restriction of ideal precision and recall

In order to deals with the semantic of alignments on one hand, and ideal precision and recall on the other hand, we first propose to use a partial closure of alignment instead of its full closure (α-consequence set). This partial closure has the advantage to be finite but in counterpart, it is defined relatively to a set of alignments. As a consequence, the ideal precision and recall can be computed, but their values depend on the set of considered alignments Λ. In a the case of evaluation campaigns, the set $\Lambda = A_{e_1} \cup ... \cup A_{e_n} \cup A_r$ will contain all correspondences provided by the participants, and the reference alignment.

Definition 12 (Bounded closure of an alignment). *The bounded closure of an alignment V given an alignment Λ ($V \subseteq \Lambda$) is defined as a set of correspondances issued from Λ which can be deduced from V.*

$$V^{+/\Lambda} = Cn(V) \cap \Lambda \tag{11}$$

The bounded closure $V^{+/\Lambda}$ of an alignment V is finite when Λ is finite (i.e. each alignment in Λ is finite). From this bounded closure definition, we provide Λ-bounded precision and recall.

Definition 13 (Λ-bounded precision measure). *Given a set of considered correspondences Λ, the precision of an alignment $A_e \subseteq \Lambda$ in comparison to a reference alignment $A_r \subseteq \Lambda$ is:*

$$P^{\Lambda}(A_e, A_r) = \frac{|A_e^{+/\Lambda} \cap A_r^{+/\Lambda}|}{|A_e^{+/\Lambda}|} \tag{12}$$

Definition 14 (Λ-bounded recall measure). *Given a set of considered correspondences Λ, the recall of an alignment $A_e \subseteq \Lambda$ in comparison to a reference alignment $A_r \subseteq \Lambda$ is:*

$$R^{\Lambda}(A_e, A_r) = \frac{|A_e^{+/\Lambda} \cap A_r^{+/\Lambda}|}{|A_r^{+/\Lambda}|} \tag{13}$$

With these measures the max-correctness, max-completeness, definiteness are verified. The semantic-identity property is also satisfied for each semantically equivalent alignments belonging to Λ (but not necessarily for the others). Still the overlapping-positiveness is not satisfied: $A_e^{+/\Lambda} \cap A_r^{+/\Lambda} = \emptyset \not\Longrightarrow Cn(A_e) \cap Cn(A_r) = \emptyset$

These measures are defined in the case of expressive alignments but they are dependent of Λ and consequently the precision and recall value are not absolute. Hence, these measures are useful for comparing a finite set of systems, but do not provide an absolute measure of precision and recall with regard to a reference alignment.

6 Conclusion

In this paper, we presented and analyzed several ontology alignment evaluation propositions. Actually, no concrete evaluation measure respects the semantic-equality and the overlapping-positiveness properties that an ideal semantic model should satisfy. More precisely, the semantic precision and recall measures cannot respect the semantic-equality due to the facts they still depend on the syntactic representation of alignments. To overcome these limitations, we first introduced alignment normalization principles which partially resolve the problem of semantic-equality. Then, we also proposed two new sets of evaluation measures. The first set of measures is built upon the generalized precision and recall framework and allows to locally consider the semantics of alignments. These measures can be seen as semantic-relaxed precision and recall. The second set of measures is proposed from an adaptation of ideal semantic measures. This adaptation makes the ideal semantic measures useable but in counterpart they do not verify the overlapping-positiveness property any more.

References

Jérôme David. *AROMA : une méthode pour la découverte d'alignements orientés entre ontologies à partir de règles d'association*. PhD thesis, Université de Nantes, 2007.

Marc Ehrig and Jérôme Euzenat. Relaxed precision and recall for ontology matching. In Benjamin Ashpole, Marc Ehrig, Jérôme Euzenat, and Heiner Stuckenschmidt, editors, *Proceedings of the Workshop on Integrating Ontologies*, volume 156, pages 25–32. CEUR-WS.org, 2005.

Jérôme Euzenat and Pavel Shvaiko. *Ontology matching*. Springer-Verlag, Heidelberg (DE), 2007.

Jérôme Euzenat. Semantic precision and recall for ontology alignment evaluation. In *Proceedings of 20th International Joint Conference on Artificial Intelligence (IJCAI 07)*, pages 248–253, Hyderabad (IN), 2007.

Jérôme Euzenat. Algebras of ontology alignment relations. In *Proceedings of the 7th International Semantic Web Conference (ISWC 08)*, pages 387–402. Springer, 2008.

Alfred Tarski. On the calculus of relations. *Journal of Symbolic Logic*, 6(3):73–89, 1941.

Cornelis Joost (Keith) van Rijsbergen. *Information Retrieval*. Butterworths, London, 2 edition, 1979.